semantically equivalent paths by associating, in a computer, paths with each other based on a similarity measure between the paths".

Table 3 in Richardson et al is not concerned with the similarity of *paths*. Instead, it is an illustration to demonstrate that the existence of the paths can be used the compute the similarity between the two *words* connected by the paths ('pen' and 'pencil'). They state in the paragraph following Table 3 that "similar words are typically connected in the MindNet by semrel paths that frequently exhibit certain pattern of relations...". Hence, Richardson uses paths to connect words, but says nothing about pairs of semantically equivalent paths or finding a similarity measure between paths. A marked up copy of section 7 in Richardson is attached that shows that Richardson deals with relations between words, and not relations between paths. Nothing in Richardson et al suggests the approach taken by the applicants. Hence, Richardson is irrelevant for all grounds of rejection. The secondary references do not supply the missing teaching.

Reconsideration and withdrawal of the rejections, and allowance of the claims, is respectfully requested.

Respectfully submitted, and certified as being faxed to the USPTO on July 5/2006

Tony Lambert

Anthony R. Lambert
Agent of Record
Registration no. 32,813
Customer no. 020212
Telephone 780-448-7326

Attachment to Response July 5th
USSN 10/826,33

Dictionaries used	LDOCE & AHD 3
Time to create (on a P2/266)	7 hours
Headwords	159,000
Definitions (N, V, ADJ)	191,000
Example sentences (N, V, ADJ)	58,000
Unique semantic relations	713,000
Inverted structures	1,047,000
Linked headwords	91,000

Table 2. Statistics on the current version of MindNet

7 Weighted paths

Inverted semrel structures facilitate the access to direct and indirect relationships between the root word of each structure, which is the headword for the MindNet entry containing it, and every other word contained in the structures. These relationships, consisting of one or more semantic relations connected together, constitute *semrel paths* between two words. For example, the semrel path between *car* and *person* in Figure 2 above is:

car←**Tobj**—*drive*—**Tsub**→*motorist*—**Hyp**→*person*.

An *extended semrel path* is a path created from subpaths in two different inverted semrel structures. For example, *car* and *truck* are not related directly by a semantic relation or by a semrel path from any single semrel structure. However, if one allows the joining of the semantic relations *car*—**Hyp**→*vehicle* and *vehicle*←**Hyp**—*truck*, each from a different semrel structure, at the word *vehicle*, the semrel path *car*—**Hyp**→*vehicle*←**Hyp**—*truck* results. Adequately constrained, extended semrel paths have proven invaluable in determining the relationship between words in MindNet that would not otherwise be connected.

Semrel paths are automatically assigned weights that reflect their salience. The weights in MindNet are based on the computation of *averaged vertex probability*, which gives preference to semantic relations occurring with middle frequency, and are described in detail in Richardson (1997). Weighting schemes with similar goals are found in work by Braden-Harder (1993) and Bookman (1994).

8 Similarity and inference

Many researchers, both in the dictionary- and corpus-based camps, have worked extensively on developing methods to identify similarity between words, since similarity determination is crucial to many word sense disambiguation and parameter-smoothing/inference procedures. However, some researchers have failed to distinguish between *substitutional* similarity and general relatedness. The similarity procedure of MindNet focuses on measuring substitutional similarity, but a function is also provided for producing clusters of generally related words.

Two general strategies have been described in the literature for identifying substitutional similarity. One is based on identifying direct, paradigmatic relations between the words, such as **Hypernym** or **Synonym**. For example, paradigmatic relations in WordNet have been used by many to determine similarity, including Li et al. (1995) and Agirre and Rigau (1996). The other strategy is based on identifying syntagmatic relations with other words that similar words have in common. Syntagmatic strategies for determining similarity have often been based on statistical analyses of large corpora that yield clusters of words occurring in similar bigram and trigram contexts (e.g., Brown et al. 1992, Yarowsky 1992), as well as in similar predicate-argument structure contexts (e.g., Grishman and Sterling 1994).

There have been a number of attempts to combine paradigmatic and syntagmatic similarity strategies (e.g., Hearst and Grefenstette 1992, Resnik 1995). However, none of these has completely integrated both syntagmatic and paradigmatic information into a single repository, as is the case with MindNet.

The MindNet similarity procedure is based on the top-ranked (by weight) semrel paths between words. For example, some of the top semrel paths in MindNet between *pen* and *pencil*, are shown below:

pen←**Means**—*draw*—**Means**→*pencil*
pen←**Means**—*write*—**Means**→*pencil*
pen—**Hyp**→*instrument*←**Hyp**—*pencil*
pen—**Hyp**→*write*—**Means**→*pencil*
pen←**Means**—*write*←**Hyp**—*pencil*

Table 3. Highly weighted semrel paths between pen and pencil

In the above example, a pattern of semrel symmetry clearly emerges in many of the paths. This observation of symmetry led to the hypothesis that similar words are typically connected in MindNet by semrel paths that frequently exhibit certain patterns of relations (exclusive of the words they actually connect), many patterns being symmetrical, but others not.

Several experiments were performed in which word pairs from a thesaurus and an anti-thesaurus (the latter containing dissimilar words) were used in a training phase to identify semrel path patterns that indicate similarity. These path patterns were then used in a testing phase to determine the substitutional similarity or dissimilarity of unseen word pairs (algorithms are described in Richardson 1997). The results, summarized in the table below, demonstrate the strength of this integrated approach, which uniquely exploits both the paradigmatic and the syntagmatic relations in MindNet.